View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Dagstuhl Research Online Publication Server

LIPIcs Leibniz International Proceedings in Informatics

The Wadge Hierarchy of Max-Regular Languages

Jérémie Cabessa[1], **Jacques Duparc**[2], **Alessandro Facchini**[2,3], **Filip Murlak**[4]

[1] Grenoble Institute of Neuroscience, Joseph Fourier University, France

[2] Faculty of Business and Economics, University of Lausanne, Switzerland

[3] LaBRI, University of Bordeaux 1, France

[4] University of Warsaw, Poland

ABSTRACT. Recently, Mikołaj Bojańczyk introduced a class of max-regular languages, an extension of regular languages of infinite words preserving many of its usual properties. This new class can be seen as a different way of generalising the notion of regularity from finite to infinite words. This paper compares regular and max-regular languages in terms of topological complexity. It is proved that up to Wadge equivalence the classes coincide. Moreover, when restricted to $\mathbf{\Delta}^0_2$-languages, the classes contain virtually the same languages. On the other hand, separating examples of arbitrary complexity exceeding $\mathbf{\Delta}^0_2$ are constructed.

Introduction

Until recently, the notion of regularity for languages of infinite words developed by Büchi [2] seemed to be universally accepted. Büchi's class has various characterisations, most notably in terms of automata and monadic second order logic, and enjoys a multitude of elegant properties, like closure by Boolean operations (including negation). Nowadays however some doubt has been cast by Mikołaj Bojańczyk [1], who presented a richer class of *max-regular languages*, arguably as much regular as Büchi's languages. This new class has a characterisation via weak monadic second-order logic with the unbounding quantifier, and a suitable automaton model with decidable emptiness. It also exhibits the usual closure properties.

In this paper we would like to shed some more light on the relations between the two classes. A typical max-regular language is defined by the property "the distance between consecutive b's is unbounded",

$$K = \{a^{n_1}ba^{n_2}ba^{n_3}\dots : \forall m\,\exists i\; n_i > m\}\,.$$

This language is not regular, but it is $\mathbf{\Pi}^0_2$-complete. In fact, as Bojańczyk notes, all max-regular languages are Boolean combinations of $\mathbf{\Sigma}^0_2$-sets, just like regular languages. Is this a coincidence, or does the similarity go further? How big is the new class? The ultimate tool for this kind of questions is the Wadge hierarchy [13, 14]. Ordering the sets based on the existence of continuous reductions (Wadge reductions) between them, the Wadge

© J. Cabessa, J. Duparc, A. Facchini, F. Murlak; licensed under Creative Commons License-NC-ND.
Foundations of Software Technology and Theoretical Computer Science (Kanpur) 2009.
Editors: Ravi Kannan and K. Narayan Kumar; pp 121–132
Leibniz International Proceedings in Informatics (LIPIcs), Schloss Dagstuhl - Leibniz-Zentrum für Informatik, Germany.
Digital Object Identifier: 10.4230/LIPIcs.FSTTCS.2009.2312

hierarchy is the most refined complexity measure in descriptive set theory. For classical regular languages, it coincides exactly with automata-based Wagner hierarchy, and is well-understood [15]. Here we investigate the Wadge hierarchy of max-regular languages.

As was shown by Finkel's work on blind counter automata [10], adding very restricted counters already makes the Wadge hierarchy much richer. Surprisingly, even though max-automata do involve counters, the Wadge hierarchy they induce actually coincides with the Wagner hierarchy. In other words, for each max-regular language, there exists a Wadge-equivalent regular language. Topologically, Bojańczyk's extension is very conservative.

On the other hand, there is an abundance of separating languages: we provide one for each level beginning from ω. This shows that the difference between the two classes spans orthogonally to the topological complexity.

Below the level ω, which corresponds exactly to the languages complete for $\mathbf{\Pi}^0_2$ or $\mathbf{\Sigma}^0_2$, the levels contain the same languages. Hence, the exemplary language K is as simple as possible: every max-regular language strictly lower than K in the Wadge hierarchy is necessarily regular.

1 Preliminaries

1.1 Languages

A set of finite words is called a *language*, and a set of infinite words an *ω-language*. Given a finite set A, called the *alphabet*, then A^*, A^+, A^ω, and A^∞ denote respectively the sets of finite words, nonempty finite words, infinite words, and finite or infinite words, all of them over the alphabet A. The empty word is denoted by ε. Given a finite word u and a finite or infinite word v, we write uv to denote the concatenation of u and v. Given $X \subseteq A^*$ and $Y \subseteq A^\infty$, the concatenation of X and Y is defined by $XY = \{xy : x \in X \text{ and } y \in Y\}$, the finite iteration of X is $X^* = \{x_1 \cdots x_n : n \geq 0 \text{ and } x_1, \ldots, x_n \in X\}$, and the infinite iteration of X is $X^\omega = \{x_0 x_1 x_2 \cdots : x_i \in X, \text{ for all } i \in \mathbb{N}\}$. Given $u \in A^*$ and $X \subseteq A^\omega$, the set $u^{-1}X$ is defined as $u^{-1}X = \{x \in A^\omega : ux \in X\}$, and X_u is $u(u^{-1}X) = uA^\omega \cap X$.

The *ω-regular languages* are exactly the ones recognised by finite Büchi, or equivalently, by finite Muller automata. We refer to [11, p.15] for further details.

Finally, for any alphabet A, the set A^ω can be equipped with the product topology of the discrete topology on A. The open sets of A^ω are thus of the form WA^ω, for some $W \subseteq A^*$.

1.2 The Wadge hierarchy

The Wadge hierarchy is a very refined topological classification of ω-languages. This classification is obtained by means of Wadge (or continuous) reduction, which is a partial ordering defined via the Wadge games [13] presented below.

Let A and B be two finite alphabets, and let $X \subseteq A^\omega$ and $Y \subseteq B^\omega$. The *Wadge game* $\mathbb{W}((A, X), (B, Y))$ is a two-player infinite game with perfect information, where player I is in charge of the subset X and player II is in charge of the subset Y. Players I and II alternately play letters from the alphabets A and B, respectively. Player I begins. Player II is allowed to skip her turn, formally denoted by the symbol "$-$", provided she plays infinitely many letters, whereas player I is not allowed to do so. After ω turns, players I and II have produced two infinite words, $\alpha \in A^\omega$ and $\beta \in B^\omega$ respectively. Player II wins $\mathbb{W}((A,X),(B,Y))$ if and only if $(\alpha \in X \Leftrightarrow \beta \in Y)$. From this point onward, the Wadge game $\mathbb{W}((A,X),(B,Y))$ will be denoted $\mathbb{W}(X,Y)$ and the alphabets involved will always be clear from the context. Along the play, the finite sequence of all previous moves of a given player is called the *current position* of this player. A *strategy* for player I is a mapping from $(B \cup \{-\})^*$ into A. A *strategy* for player II is a mapping from A^+ into $B \cup \{-\}$. A strategy is *winning* if the player following it must necessarily win, no matter what his opponent plays.

The *Wadge reduction* is defined via the Wadge game as follows: a set X is said to be *Wadge reducible* to Y, denoted by $X \leq_W Y$, if and only if player II has a winning strategy in $\mathbb{W}(X,Y)$. This relation $\leq_W$ is reflexive and transitive. The corresponding equivalence relation and strict reduction are defined by $X \equiv_W Y$ if and only if both $X \leq_W Y$ and $Y \leq_W X$ hold, and $X <_W Y$ if and only if $X \leq_W Y$ and $X \not\equiv_W Y$. In addition, the sets X and Y are said to be Wadge incomparable, denoted as $X \perp_W Y$, if and only if both $X \not\leq_W Y$ and $Y \not\leq_W X$. Besides, a set $X \subseteq A^\omega$ is called *self-dual* if $X \equiv_W X^c$, and *non-self-dual* if $X \not\equiv_W X^c$.

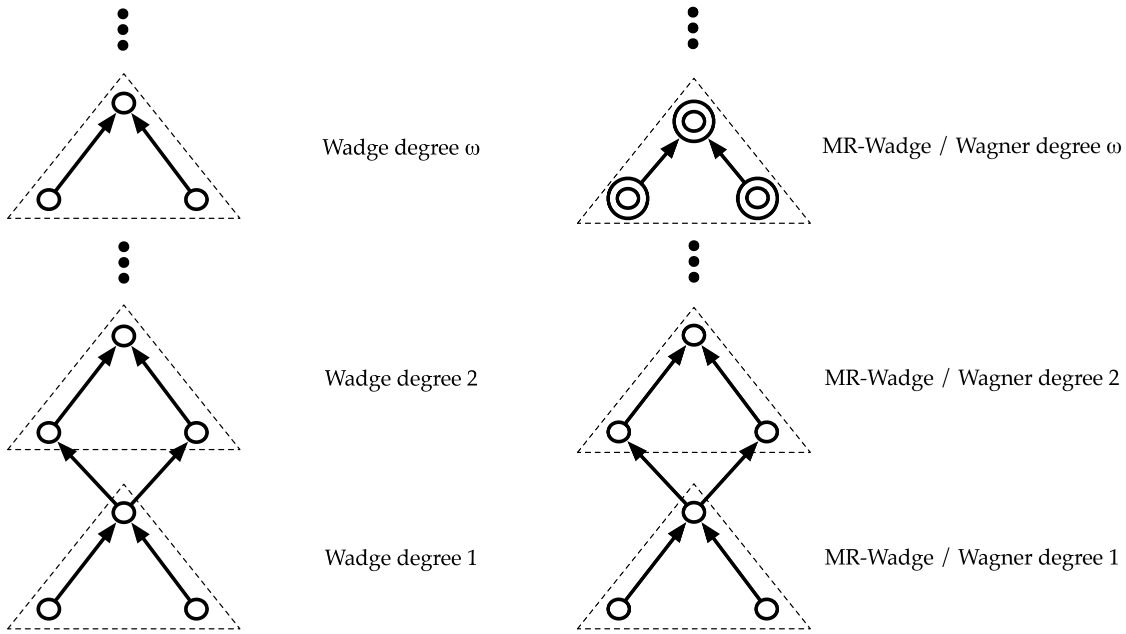
Let us point out that Wadge games were designed so that the Wadge reduction correspond precisely to the continuous reduction. Indeed, it holds that $X \leq_W Y$ if and only if there exists a continuous function $f : A^\omega \to B^\omega$ such that $f^{-1}(Y) = X$ [13].

The Wadge hierarchy consists of the collection of all ω-languages ordered by the Wadge reduction, and *the Borel Wadge hierarchy* is the restriction of the Wadge hierarchy to Borel ω-languages. As a consequence of Martin's Borel determinacy theorem, for any two Borel ω-languages X and Y, there exists a winning strategy for one of the players in $\mathbb{W}(X,Y)$. This key property induces the following strong consequences on the Borel Wadge hierarchy. First, the $\leq_W$-antichains have length at most 2, and the only incomparable ω-languages are, up to Wadge equivalence, of the form X and X^c, for X non-self-dual. Furthermore, the Wadge reduction is well-founded on Borel sets, meaning that there is no infinite strictly descending sequence of Borel ω-languages $X_0 >_W X_1 >_W X_2 >_W \ldots$. These results ensure that, up to complementation and Wadge equivalence, the Borel Wadge hierarchy is actually a well ordering.

Therefore, there exist a unique ordinal, called the *height* of the Borel Wadge hierarchy, and a mapping d_W from the Borel Wadge hierarchy onto its height, called the *Wadge degree*, such that $d_W(X) < d_W(Y)$ if and only if $X <_W Y$, and $d_W(X) = d_W(Y)$ if and only if either $X \equiv_W Y$ or $X \equiv_W Y^c$, for every Borel ω-languages X and Y. Actually, it is usually convenient to consider another definition of the Wadge degree which makes the non-self dual sets and the first self dual ones that strictly reduce these latter always share the same degree, namely:

$$d_W(X) = \begin{cases} 1 & \text{if } X = \emptyset \text{ or } X = \emptyset^c, \\ \sup\{d_W(Y) + 1 : Y \text{ n.s.d. and } Y <_W X\} & \text{if } X \text{ is non-self-dual}, \\ \sup\{d_W(Y) : Y \text{ n.s.d. and } Y <_W X\} & \text{if } X \text{ is self-dual}. \end{cases}$$

Furthermore, it can be proved that the Borel Wadge hierarchy actually consists of an alternating succession of non-self-dual and self-dual sets with non-self-dual pairs at each limit level (provided finite alphabets are considered) [7, 13, 14]. Therefore, for any ordinal α below the height of the Borel Wadge hierarchy, there exist exactly three Wadge classes of degree α, namely two non-self-dual and one self-dual located precisely just one level above, as illustrated in Figure 1(a).



(a) The Wadge hierarchy: circles represent Wadge-equivalence classes and arrows stand for the strict Wadge reduction between those. The non-self dual sets and the self dual ones located just one level above share the same Wadge degree.

(b) The MR-Wadge and the Wagner hierarchy. On finite levels the classes coincide; above, MR-Wadge classes properly extend corresponding Wagner classes.

Figure 1: The hierarchies

The three Wadge classes are very closely related. In fact, any set $X \subseteq A^\omega$ that is complete for some Wadge class of degree α gives rise to two other sets $Y, Z \subseteq A^\omega$ that are respectively complete for the two remaining Wadge classes of same degree α. More precisely, if one starts with X self-dual such that $d_W(X) = \alpha$, then we know that there exists $u \in A^*$ such that $Y = u^{-1}X$ is non-self-dual and $d_W(Y) = \alpha$. It directly follows that $Z = (u^{-1}X)^c$ is also non-self-dual and $d_W(Z) = \alpha$. On the other hand, if one starts with X non-self-dual and $d_W(X) = \alpha$, then $Y = X^c$ is also non-self-dual, Wadge incomparable with X, and $d_W(Y) = \alpha$. Moreover, for any $a \in A$, the set $Z = aX \cup (A \setminus \{a\})X^c$ is self-dual with $d_W(Z) = \alpha$. All these results are folklore and can be found for instance in [7]. In the sequel we will also use the fact that the constructions above preserve regularity and max-regularity.

In this paper we are working only with the sets from $\mathrm{BC}(\boldsymbol{\Sigma}^0_2)$, the class of Boolean combinations of $\boldsymbol{\Sigma}^0_2$ sets, but in fact we need to go quite deep into the structure of the Wadge hierarchy in order to obtain the promised results. The proofs of all the facts we state below can be found in [7].

Let us start with the relation between the Borel classes and the Wadge degrees. The nth level of the Borel hierarchy corresponds to the Wadge degree "a tower of ω_1's of the height $n-1$". In particular, a language complete for $\boldsymbol{\Sigma}^0_2$ or $\boldsymbol{\Pi}^0_2$ has degree ω_1. This already shows how drastically the Borel Wadge hierarchy refines the Borel hierarchy! When we move to combinations of $\boldsymbol{\Sigma}^0_2$ sets, we get exactly the Wadge degrees strictly below ω_1^ω.

Important milestones on the way from ω_1 to ω_1^ω are the so-called *initialisable* sets. They are defined as those sets X, for which player II has a winning strategy in the II-imposed Wadge game $\mathbb{W}(X, X)$ where player I is allowed at any moment, but only once, to erase everything he has played before and start anew.

Let us remark that initialisable sets generalise prefix-independent sets, i.e., sets satisfying condition $u^{-1}X = X$ for all finite words u. Indeed, the winning strategy for player II in the corresponding game amounts to copying the letters played by player I, even after player I decides to erase everything and start again: the part of player II's word played before player I erased his word will not influence the outcome. Roughly speaking, initialisability is prefix-independence up to Wadge-equivalence.

Initialisable sets within $\mathrm{BC}(\mathbf{\Sigma}_2^0)$ are exactly those with Wadge degrees ω_1^n for some natural number n. Clearly, the empty set and the whole space are prefix-independent, and so initialisable. So is the well-known $\mathbf{\Pi}_2^0$-complete set $(1^*2)^\omega$. In fact, the parity languages with $n+1$ ranks correspond exactly to the degree ω_1^n. Showing that no other degree below ω_1^ω is initialisable requires a lot of technical effort. We refer the reader to [7] for the proof.

Let us finish this quick peek into the internal structure of $\mathrm{BC}(\mathbf{\Sigma}_2^0)$ with a fact that shows how simpler sets are hidden inside more complex ones. As already stated, $\mathrm{BC}(\mathbf{\Sigma}_2^0)$ sets have degrees strictly below ω_1^ω. Hence, if $X \subseteq A^*$ is $\mathrm{BC}(\mathbf{\Sigma}_2^0)$, its Wadge degree can be written in the Cantor normal form of base ω_1 as $d_W(X) = \omega_1^{n_k} \cdot p_k + \cdots + \omega_1^{n_0} \cdot p_0$, for some $k > 0$, some $\omega > n_k > \ldots > n_0 \geq 0$, and some $0 < p_i < \omega_1$ for all $0 \leq i \leq k$. Assume that one of the coefficients, say p_j, is not finite, i.e., $p_j \geq \omega$. Then for each $m > 0$ there exists a word $u \in A^*$ such that $d_W(X_u) = \omega_1^{n_k} \cdot p_k + \cdots + \omega_1^{n_j} \cdot m$. This fact is a special case of a more general result [8, Lemmas 33 and 39]. The following lemma follows easily.

LEMMA 1. *Let $X \subseteq A^*$ be a $\mathrm{BC}(\mathbf{\Sigma}_2^0)$ set such that the family $\{X_u \colon u \in A^*\}$ is finite up to Wadge equivalence. Then*

$$d_W(X) = \omega_1^{n_k} \cdot p_k + \cdots + \omega_1^{n_0} \cdot p_0\,,$$

for some $k > 0$, some $\omega > n_k > \ldots > n_0 \geq 0$, and some $0 < p_i < \omega$ for all $0 \leq i \leq k$.

1.3 The Wagner hierarchy

In 1979, Klaus Wagner described a classification of ω-regular sets in terms of the graph-theoretical structure automata known as the *the Wagner hierarchy* [15]. This hierarchy is a decidable pre-well-ordering of width 2 and height ω^ω. The Wagner degree of any given ω-regular language can be effectively computed by analysing the graph of a Muller automaton accepting this language [16].

In 1986, Simonnet proved that the Wagner hierarchy corresponds precisely to the restriction of the Wadge hierarchy to ω-regular languages. In our further explanations the following notion will be convenient. We say that a Wadge class is *inhabited* by a language if the language is complete for the Wadge class. In these terms, ω-regular languages inhabit exactly all Wadge classes with Wadge degrees of the form $\omega_1^{n_k} \cdot p_k + \cdots + \omega_1^{n_0} \cdot p_0$, where $\omega > n_k > \ldots > n_0 \geq 0$ and $0 < p_i < \omega$ for all $0 \leq i \leq k$. In addition, it can be shown that the Wagner reduction, which already coincides with the Wadge reduction, can also be defined in terms of automata [11, Thm. 5.2, p. 209]. Similarly to the Wadge degree, the Wagner degree of an ω-regular language L can thus be defined as follows:

$$d_{\omega R}(L) = \begin{cases} 1 & \text{if } L = \emptyset \text{ or } L = \emptyset^c, \\ \sup\{d_{\omega R}(K) + 1 : K \text{ n.s.d. and } K <_W L\} & \text{if } L \text{ is non-self-dual}, \\ \sup\{d_{\omega R}(K) : K \text{ n.s.d. and } K <_W L\} & \text{if } L \text{ is self-dual.} \end{cases}$$

In consequence, the Wagner and the Wadge degrees of ω-regular languages are related as follows: for any ω-regular language L, if

$$d_{\omega R}(L) = \omega^{n_k} \cdot p_k + \cdots + \omega^{n_0} \cdot p_0\,,$$

for some $\omega > n_k > \ldots > n_0 \geq 0$ and $0 < p_i < \omega$ for all $0 \leq i \leq k$, then

$$d_W(L) = \omega_1^{n_k} \cdot p_k + \cdots + \omega_1^{n_0} \cdot p_0\,.$$

The Wagner hierarchy has been extensively investigated. Its complete set theoretical description in terms of Boolean expressions was given by Selivanov [12], and its algebraic counterpart was studied by various authors [3, 4, 5, 6, 9].

2 Max-regular languages

In [1], Bojańczyk introduces a new class of languages of infinite words called *max-regular languages*. This class is a proper extension of the class of ω-regular languages. It has two equivalent descriptions, one in terms of automata (max-automata), and the other in terms of logic (weak MSO with the unbounding quantifier). Here, we briefly recall the automata-theoretic one.

DEFINITION 2. *A max-automaton is a tuple $\mathcal{A} = (Q, A, \Gamma, q_0, E, \mathcal{T})$, where Q is a finite set of states, A a finite input alphabet, Γ a finite set of counters, q_0 an initial state, $\mathcal{T} \subseteq \mathcal{P}(\Gamma)$ is a specified collection of subsets of Γ, and $E \subseteq Q \times A \times Q \times (\bigcup_{c,c' \in \Gamma}\{inc_c, res_c, out_c, max_{c,c'}\})^*$ is a finite set of transitions, which, given a current state q and input letter a specifies a changing state and a sequence of counter operations. The operations inc_c, res_c, out_c, and $max_{c,c'}$ respectively mean set $c := c + 1$, set $c := 0$, output the current value of c, and set $c := \max(c, c')$.*

As usual, a deterministic max-automaton is defined by requiring the transition set E to be the graph of a partial function from $Q \times A$ into $Q \times (\bigcup_{c,c' \in \Gamma}\{inc_c, res_c, out_c, max_{c,c'}\})^*$.

For any counter $c \in \Gamma$ and any finite sequence of counter operations $o_0, \ldots, o_i$, the value of counter c after the successive performing of these operations will be denoted by $c(o_1 \cdots o_i)$.

A run of $\mathcal{A}$ is a sequence of consecutive transitions. Given an infinite run ρ, the infinite output sequence of counter c during ρ is denoted by ρ_c. An infinite word x is *accepted* by $\mathcal{A}$ if it admits a run ρ such that $\{c \in \Gamma : \rho_c \text{ is unbounded}\} \in \mathcal{T}$. In other words, the accepting conditions of max-automata are Boolean combinations of clauses of the form "the sequence ρ_c is bounded".

The set of infinite words accepted by $\mathcal{A}$ is the *language recognised by* $\mathcal{A}$ and is denoted by $L(\mathcal{A})$. An ω-language is called *max-regular* if it is recognised by a *deterministic* max-automaton.

Note that, as for Muller automata, up to adding a sink state together with the appropriate transitions and counter operations, we may assume without loss of generality that every deterministic max-automaton is complete. Hence, for any finite or infinite word, there exists exactly one corresponding finite or infinite run labelled by this word. From this point onwards, every max-automaton will be assumed to be deterministic and complete.

The following fact is taken from [1]. We sketch the proof for the sake of completeness.

LEMMA 3. *The class of max-languages is a proper extension of the class of ω-regular languages.*

PROOF. The language $K = \{a^{n_1}ba^{n_2}ba^{n_3}\ldots : \forall m\, \exists i\, n_i > m\}$ mentioned in the introduction separates the classes. Let us concentrate on showing that every ω-regular language is max-regular.

Let L be an ω-regular language, and let $\mathcal{A} = (Q, A, q_0, \delta, \mathcal{T})$ be a deterministic Muller automaton recognising it. We build a deterministic max-automaton $\mathcal{A}'$ recognising this same language. The automaton $\mathcal{A}' = (Q', A, \Gamma, q_0', \delta', \mathcal{T}')$ is obtained by associating a counter c_q with each state q of $\mathcal{A}$ and by simulating the visit of each state of $\mathcal{A}$ by incrementing and outputting the corresponding counter of $\mathcal{A}'$. More precisely, we set $Q' = Q$, $\Gamma = \{c_q : q \in Q\}$, $q_0' = q_0$, $\delta' = \{(q, a, q', (inc_{c_q'}, out_{c_q'})) : (q, a, q') \in \delta\}$, and $\mathcal{T}' = \{\{c_{q_1}, \ldots, c_{q_n}\} : \{q_1, \ldots, q_n\} \in \mathcal{T}\}$. In this way, a state of $\mathcal{A}$ is visited infinitely often iff the output sequence of its corresponding counter in $\mathcal{A}'$ is unbounded. The definition of $\mathcal{T}'$ then ensures that $\mathcal{A}$ and $\mathcal{A}'$ recognise the same ω-language. $\square$

We now prove that if two infinite words induce converging runs, they are either both accepted or both rejected. This technical result will be very useful in the sequel. For finite words u and v we write $u \sim_{\mathcal{A}} v$ iff $\mathcal{A}$'s runs on u and v end in the same state.

LEMMA 4. *Let $\mathcal{A}$ be a deterministic max-automaton, and let u and v such that $u \sim_{\mathcal{A}} v$. Then $u^{-1}L(\mathcal{A}) = v^{-1}L(\mathcal{A})$.*

PROOF. Let A be the input alphabet of the automaton $\mathcal{A}$, and let $x = x_0x_1x_2\cdots$ be some infinite word of A^ω. Let also $\rho = \rho_0\rho_1\rho_2\cdots$ and $\rho' = \rho_0'\rho_1'\rho_2'\cdots$ be the two infinite runs of $\mathcal{A}$ labelled by ux and vx, respectively, and let $o_0o_1o_2\cdots$ and $o_0'o_1'o_2'\cdots$ be the two corresponding infinite sequences of counter operations performed during these respective runs. Since $u \sim_{\mathcal{A}} v$, there exist two integers m' and n' such that $\rho_{m'+i} = \rho'_{n'+i}$ for all $i \geq 0$, thus there also exist two integers m and n such that $o_{m+i} = o'_{n+i}$ for all $i \geq 0$. Now let $k = \max_{c\in\Gamma} |c(o_0\cdots o_m) - c(o_0'\cdots o_n')|$. We prove by induction on $i \in \mathbb{N}$ that the relation $|c(o_0\cdots o_{m+i}) - c(o_0'\cdots o'_{n+i})| \leq k$ holds for all $c \in \Gamma$.

By definition of k, the claim holds for $i = 0$. Now let $i > 0$, and assume that for all $j \leq i$, the inequality $|c(o_0\cdots o_{m+j}) - c(o_0'\cdots o'_{n+j})| \leq k$ is true for all $c \in \Gamma$. Let $c \in \Gamma$, and consider the counter operation $o_{m+i+1} = o'_{n+i+1}$. We discuss the nature of this operation.

(1) If $o_{m+i+1} = o'_{n+i+1} = res_c$, then $|c(o_0\cdots o_{m+i+1}) - c(o_0'\cdots o'_{n+i+1})| = 0 \leq k$.

(2) If $o_{m+i+1} = o'_{n+i+1}$ is either inc_c or out_c, then by the induction hypothesis, it follows that $|c(o_0\cdots o_{m+i+1}) - c(o_0'\cdots o'_{n+i+1})| = |c(o_0\cdots o_{m+i}) - c(o_0'\cdots o'_{n+i})| \leq k$.

(3) If $o_{m+i+1} = o'_{n+i+1}$ concerns another counter than c, then by the induction hypothesis $|c(o_0 \cdots o_{m+i+1}) - c(o'_0 \cdots o'_{n+i+1})| = |c(o_0 \cdots o_{m+i}) - c(o'_0 \cdots o'_{n+i})| \leq k$.

(4) If $o_{m+i+1} = o'_{n+i+1} = max_{c,d}$, for some $d \in \Gamma$, four different cases need to be considered:

(a) If $c(o_0 \cdots o_{m+i}) \leq d(o_0 \cdots o_{m+i})$ and $c(o'_0 \cdots o'_{n+i}) \leq d(o'_0 \cdots o'_{n+i})$, it follows that $c(o_0 \cdots o_{m+i+1}) := d(o_0 \cdots o_{m+i})$ and $c(o'_0 \cdots o'_{n+i+1}) := d(o'_0 \cdots o'_{n+i})$. Therefore by the induction hypothesis $|c(o_0 \cdots o_{m+i+1}) - c(o'_0 \cdots o'_{n+i+1})| = |d(o_0 \cdots o_{m+i}) - d(o'_0 \cdots o'_{n+i})| \leq k$.

(b) The case $c(o_0 \cdots o_{m+i}) \geq d(o_0 \cdots o_{m+i})$ and $c(o'_0 \cdots o'_{n+i}) \geq d(o'_0 \cdots o'_{n+i})$ is symmetric.

(c) If $c(o_0 \cdots o_{m+i}) \leq d(o_0 \cdots o_{m+i})$ but $c(o'_0 \cdots o'_{n+i}) \geq d(o'_0 \cdots o'_{n+i})$, it follows that $c(o_0 \cdots o_{m+i+1}) := d(o_0 \cdots o_{m+i})$ and $c(o'_0 \cdots o'_{n+i+1}) := c(o'_0 \cdots o'_{n+i})$. Thence $|c(o_0 \cdots o_{m+i+1}) - c(o'_0 \cdots o'_{n+i+1})| = |d(o_0 \cdots o_{m+i}) - c(o'_0 \cdots o'_{n+i})|$. Now the two following cases need to be distinguished:

i. If $c(o'_0 \cdots o'_{n+i}) \leq d(o_0 \cdots o_{m+i})$, thence $|d(o_0 \cdots o_{m+i}) - c(o'_0 \cdots o'_{n+i})| = d(o_0 \cdots o_{m+i}) - c(o'_0 \cdots o'_{n+i}) \leq d(o_0 \cdots o_{m+i}) - d(o'_0 \cdots o'_{n+i}) \leq k$.

ii. If $c(o'_0 \cdots o'_{n+i}) \geq d(o_0 \cdots o_{m+i})$, thence $|d(o_0 \cdots o_{m+i}) - c(o'_0 \cdots o'_{n+i})| = c(o'_0 \cdots o'_{n+i}) - d(o_0 \cdots o_{m+i}) \leq c(o'_0 \cdots o'_{n+i}) - c(o'_0 \cdots o'_{n+i}) \leq k$.

(d) The case $c(o_0 \cdots o_{m+i}) \geq d(o_0 \cdots o_{m+i})$ but $c(o'_0 \cdots o'_{n+i}) \leq d(o'_0 \cdots o'_{n+i})$ is symmetric.

Now since $|c(o_0 \cdots o_{m+i}) - c(o'_0 \cdots o'_{n+i})| \leq k$ for all $i \geq 0$ and all $c \in \Gamma$, it follows that, for all $c \in \Gamma$, the output sequence ρ_c is bounded iff ρ'_c is also bounded. Therefore $ux \in L(\mathcal{A})$ iff $vx \in L(\mathcal{A})$ for all $x \in A^\omega$, or in other words, $u^{-1}L(\mathcal{A}) = v^{-1}L(\mathcal{A})$. □

3 The Wadge hierarchy of max-regular languages

The collection of all max-regular languages ordered by the Wadge reduction will be called the *MR-Wadge hierarchy*. The present section provides a description of this hierarchy. We prove that, although the class of max-regular languages properly extends the class of ω-regular languages, the MR-Wadge hierarchy and the Wagner hierarchy are equal up to Wadge equivalence.

THEOREM 5. *Max-regular languages inhabit exactly those self-dual and non-self-dual classes, which have the Wadge degree of the form*

$$\omega_1^{n_k} \cdot p_k + \cdots + \omega_1^{n_0} \cdot p_0$$

with $k > 0$, $\omega > n_k > \ldots > n_0 \geq 0$, *and* $0 < p_i < \omega$ *for all* $0 \leq i \leq k$.

In particular, the MR-Wadge hierarchy is a pre-well-ordering of width 2 *and height* ω^ω.

PROOF. Let α be an ordinal with Cantor normal form $\alpha = \omega_1^{n_k} \cdot p_k + \cdots + \omega_1^{n_0} \cdot p_0$, for some $k > 0$, some $\omega > n_k > \ldots > n_0 \geq 0$ and some $0 < p_i < \omega$ for all $0 \leq i \leq k$. In the Wagner hierarchy, there exist two ω-regular languages L and L' such that L is self-dual, L' is non-self dual, and $d_W(L) = d_W(L') = \alpha$. Lemma 3 guarantees that L and L' are also max-regular.

It remains to prove that no other Wadge class is inhabited by a max-regular language. Let L be a max-regular language over the alphabet A. The language L is recognised by a finite state max-automaton, so from Lemma 4 it follows that the family $\{u^{-1}L : u \in A^*\}$ is finite. But then, up to Wadge equivalence, $\{L_u : u \in A^*\}$ is finite and the claim follows by Lemma 1. □

More precisely, the MR-Wadge hierarchy consists of an alternating succession of non-self-dual and self-dual Wadge classes with non-self-dual pairs at each limit level. The MR degree of a max-regular language L is now defined as

$$d_{MR}(L) = \begin{cases} 1 & \text{if } L = \emptyset \text{ or } L = \emptyset^c, \\ \sup\{d_{MR}(K) + 1 : K \text{ n.s.d. and } K <_W L\} & \text{if } L \text{ is non-self-dual}, \\ \sup\{d_{MR}(K) : K \text{ n.s.d. and } K <_W L\} & \text{if } L \text{ is self-dual}. \end{cases}$$

Once again, this definition of the MR degree ensures that the non-self dual languages and the self dual ones located just one level above in the MR-Wadge hierarchy always share the same degree. Therefore, the MR-Wadge and the Wadge degrees of max-regular languages are related as follows: for any max-regular languages L, if $d_{MR}(L) = \omega^{n_k} \cdot p_k + \cdots + \omega^{n_0} \cdot p_0$, for some $\omega > n_k > \ldots > n_0 \geq 0$ and $0 < p_i < \omega$ for all $0 \leq i \leq k$, then $d_W(L) = \omega_1^{n_k} \cdot p_k + \cdots + \omega_1^{n_0} \cdot p_0$.

4 The MR-Wadge and the Wagner hierarchies

We now provide a detailed comparison of the MR-Wadge and the Wagner hierarchies. In the previous section we have seen that the MR-Wadge and the Wagner hierarchies inhabit exactly the same Wadge classes.

THEOREM 6. *The MR-Wadge and the Wagner hierarchy are equal (up to Wadge equivalence).*

The following two results prove that the ω first classes of the MR-Wadge and the Wagner hierarchies contain exactly the same ω-languages, whereas every other MR-Wadge class is a proper extension of its Wagner counterpart (see Fig. 1(b)).

PROPOSITION 7. *For every natural number n the following conditions are equivalent:*
(1) L is ω-regular and $d_{\omega R}(L) = n$.
(2) L is max-regular and $d_{MR}(L) = n$.

PROOF. Let us first see that (1) implies (2). Let L be ω-regular with $d_{\omega R}(L) = n$. Then L is also max-regular. Moreover, the structure of the Wagner hierarchy ensures that $d_W(L) = n$. Hence, by Theorem 5, $d_{MR}(L) = n$.

Now, let us prove that (2) implies (1). Take a max-regular language L with $d_{MR}(L) = n$. We first show that L is ω-regular. Let $\mathcal{A} = (Q, A, \Gamma, q_0, \delta, \mathcal{T})$ be a max-automaton that recognises L. Let $\mathcal{C}_1, \ldots, \mathcal{C}_p$ be all (maximal) strongly connected components (s.c.c.) of the graph of the automaton $\mathcal{A}$. Given any infinite word x, we denote $\mathrm{scc}(x)$ the unique s.c.c. that contains all states visited infinitely often while reading x. In other words, $\mathrm{scc}(x)$ is the s.c.c. inside which the reading of the terminal part of x takes place. Consider the following equivalence relation between infinite words: $x \approx y$ iff $\mathrm{scc}(x) = \mathrm{scc}(y)$. We claim that $x \approx y$ implies that $(x \in L \Leftrightarrow y \in L)$. Towards a contradiction, assume that there exist $x \in L$ and

$y \notin L$ with $x \approx y$. Let $\mathrm{scc}(x) = \mathrm{scc}(y) = \mathcal{C}_i$ and let $u, v \in A^*$ be the shortest prefixes of x and y respectively such that there exist respectively $q_u, q_v \in \mathcal{C}_i$ with $q_0 \xrightarrow{u} q_u$ and $q_0 \xrightarrow{v} q_v$. Let x', y' be such that $x = ux'$ and $y = vy'$. Since $\mathcal{C}_i$ is a s.c.c., there exists a finite word w such that $q_u \xrightarrow{w} q_v$. Consider $Z = \{z \in uA^\omega : \mathrm{scc}(z) = \mathcal{C}_i\}$. We next prove the following facts:

(1) $Z \cap L$ is initialisable,
(2) both $\emptyset \leq_W Z \cap L$ and $\emptyset^c \leq_W Z \cap L$ hold,
(3) $Z \cap L \leq_W L$.

(1) Consider the II-imposed game $\mathbb{W}(Z \cap L, Z \cap L)$ where I may only once erase his play and start anew. We will provide a winning strategy for player II that guarantees that she *always remains inside* Z. As long as player I stays inside Z, player II should copy his actions. If player I exits Z, player II should play a finite word that reaches q_v, and then to play y'. If player I decides to erase everything he has played since the beginning, then player II can still catch up by playing any finite word that leads her back to q_u, and start copying again I's play, from the moment when I reaches q_u. If player I exits Z again, II should proceed like before. By Lemma 4 this provides a winning strategy. (2) $\emptyset \leq_W Z \cap L$ and $\emptyset^c \leq_W Z \cap L$ hold because playing $x = uwy'$ or ux', respectively, is winning for II in the corresponding Wadge games. (3) A winning strategy for player II in $\mathbb{W}(Z \cap L, L)$ amounts to copying player I's moves, as long as he stays in Z. If player I exits Z, player II should play a word reaching q_v (this is always possible, since so far player II has stayed inside Z) and then play y'.

Since $Z \cap L$ is a Boolean combination of $\mathbf{\Sigma}^0_2$ sets, by a result from [7], condition (1) yields $d_W(Z \cap L) = \omega_1^n$ for some natural n. Condition (2) ensures that $n > 0$, hence $d_W(Z \cap L) \geq \omega_1$. Finally, condition (3) implies that $d_W(L) \geq \omega_1$, but this is a contradiction. Hence, the claim holds.

Consider $\mathcal{A}' = (Q, A, q_0, \delta', F)$, the deterministic finite automaton with Büchi acceptance conditions where δ' is just δ with the operations on counters removed, and F is the set of states q for which there exists an infinite word $x \in L$ such that $q \in \mathrm{scc}(x)$. Then $\mathcal{A}'$ recognises L, which shows that L is ω-regular. Theorem 6, guarantees that $d_{MR}(L) = d_{\omega R}(L) = d_W(L) = n$. ☐

Before we move to the proof of our last result, let us show that the language

$$K = \{a^{n_1}ba^{n_2}ba^{n_3}b \cdots : \forall m\ \exists i\ n_i > m\}$$

is $\mathbf{\Pi}^0_2$-complete, as stated in the introduction. It is very easy to see that it is Wadge equivalent to the $\mathbf{\Pi}^0_2$-complete $L' = (a^*b)^\omega$. Indeed, player II has a winning strategy in the game $\mathbb{W}(L, L')$: every time player I produces a sequence of consecutive a's that is strictly longer than all previous ones, Player II should play a b. Otherwise, player II should play an a. Conversely, player II also has a winning strategy in the game $\mathbb{W}(L', L)$: every time player I plays a b, player II should play a sequence of consecutive a's that is strictly longer than all previously played, followed by b. Otherwise, she should play b alone.

PROPOSITION 8. *Let $\alpha = \omega_1^{n_k} \cdot p_k + \cdots + \omega_1^{n_0} \cdot p_0 \geq \omega_1$, where $\omega > n_k > \ldots > n_0 \geq 0$ and $0 < p_i < \omega$ for all $0 \leq i \leq k$. Then there exist max-regular languages L and L' such that L is self-dual, L' is non-self-dual, $d_W(L) = d_W(L') = \alpha$, and both L and L' are not ω-regular.*

PROOF. Without loss of generality we may assume that $A = \{a, b\}$. We first prove the existence of appropriate non-self-dual languages over A. If $\alpha = \omega_1$, then consider the language K above. It is $\mathbf{\Pi}^0_2$-complete, which means that $d_W(K) = \omega_1 = \alpha$, as mentioned in Sect. 1.2. Now if $\alpha = \omega_1^{n_k} \cdot p_k + \cdots + \omega_1^{n_0} \cdot p_0 > \omega_1$, then there exists a non-self-dual ω-regular language $M \subseteq A^\omega$ such that $d_W(M) = \alpha$. Let $L = aM \cup bK$. The language L is non-self-dual and satisfies $L \equiv_W M$. Thus $d_W(L) = d_W(M) = \alpha$. In addition, since both M and K are max-regular, so is L. Finally, L is not ω-regular, for if it were so, then $b^{-1}L = K$ would also be ω-regular – a contradiction.

From the existence of an appropriate non-self-dual language, we deduce the existence of an appropriate self-dual language over A. Let $L \subseteq A^\omega$ be a non-self-dual max-regular language such that both $d_W(L) = \alpha$ and L is not ω-regular. Take $L' = aL \cup bL^c$. Then L' is also max-regular. Moreover, as mentioned in Sect. 1.2, L' is self-dual and $d_W(L') = \alpha$. Finally, L' is not ω-regular, for if it were so, the language $a^{-1}L' = L$ would also be ω-regular – a contradiction. $\square$

Conclusion

We have given a precise comparison of the Wadge hierarchies for regular and max-regular languages. As the hierarchies coincide, Bojańczyk's extension does not increase the topological complexity. It does provide more variety though, as witnessed by the plethora of separating examples.

The results of this paper give a complete description of the Wadge hierarchy of max-regular languages. Alas, the description is not effective (unlike [10, 15]). What is missing is an algorithm to decide the Wadge degree of a given language. From the proof of Proposition 7 one could extract a partial decidability result. Using decidability of emptiness for max-automata, one can check if there are two words $x \in L(A)$ and $y \notin L(A)$, such that the runs on both of them are finally trapped in the same strongly connected component of A, thus deciding if $L(A)$ is at least on the level ω or not. If not, one can construct effectively an equivalent automaton without counters, and use the Wagner's characterisation to compute the exact degree. Obtaining decidability of higher levels would probably require much deeper analysis of the loop structure within strongly connected components. We point this out as a promising line of investigation.

As for the technical side of the paper, we would like to highlight the method used to prove that no other Wadge degrees are realised by max-regular languages (Theorem 5). Here, the argument relies on the fact that the family $\{w^{-1}L \colon w \in A^*\}$ is finite up to Wadge equivalence. A more involved version of this method, based on a generalisation of Lemma 1, has been successfully applied to deterministic push-down automata [8]. We believe that this technique can be useful for other models of computation as well.

Acknowledgement

The fourth author was supported by the Polish government grant no. N206 008 32/0810. Part of this work was carried out when the fourth author was postdoc at the University of Edinburgh, UK. The suggestions of the anonymous referees were very helpful in the preparation of the final version of this paper.

References

[1] M. Bojańczyk. Weak MSO with the unbounding quantifier. STACS 2009: 159–170.

[2] J. R. Büchi. On a decision method in restricted second order arithmetic. *Proc. 1960 Int. Congr. for Logic, Methodolog, and Philosophy of Science*, Stanford Univ. Press, 1962, 1–11.

[3] J. Cabessa, J. Duparc. A game theoretical approach to the algebraic counterpart of the Wagner hierarchy: Part i. *RAIRO-Theor. Inf. Appl.* 43(3): 443–461 (2009).

[4] J. Cabessa, J. Duparc. A game theoretical approach to the algebraic counterpart of the Wagner hierarchy: Part ii. *RAIRO-Theor. Inf. Appl.* 43(3): 463–515 (2009).

[5] O. Carton, D. Perrin. Chains and superchains for ω-rational sets, automata and semigroups. *Int. J. Algebra Comput.*, 7(6): 673–695 (1997).

[6] O. Carton, D. Perrin. The Wagner hierarchy. *Int. J. Algebra Comput.* 9(5):597–620 (1999).

[7] J. Duparc. Wadge hierarchy and Veblen hierarchy. Part i: Borel sets of finite rank. *J. Symb. Log.* 66(1): 56–86 (2001).

[8] J. Duparc. A hierarchy of deterministic context-free ω-languages. *Theor. Comput. Sci.* 290(3): 1253–1300 (2003).

[9] J. Duparc, M. Riss. The missing link for ω-rational sets, automata, and semigroups. *Int. J. Algebra Comput.* 16(1): 161–185 (2006).

[10] O. Finkel. An Effective Extension of the Wagner Hierarchy to Blind Counter Automata. CSL 2001: 369-383

[11] D. Perrin, J.-É. Pin. *Infinite words*. Volume 141 of *Pure and Applied Mathematics*, Elsevier, 2004.

[12] V. Selivanov. Fine hierarchy of regular ω-languages. *Theor. Comput. Sci.* 191(1-2): 37–59 (1998).

[13] W. W. Wadge. Degrees of complexity of subsets of the Baire space. *Notices A.M.S.*, pp. A714–A715, 1972.

[14] W. W. Wadge. *Reducibility and determinateness on the Baire space*. PhD thesis, University of California, Berkeley (1983).

[15] K. Wagner. On ω-regular sets. *Inform. and Control* 43(2): 123–177 (1979).

[16] T. Wilke, H. Yoo. Computing the Wadge degree, the Lifshitz degree, and the Rabin index of a regular language of infinite words in polynomial time. TAPSOFT 1995: 288–302.

This work is licensed under the Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 License.